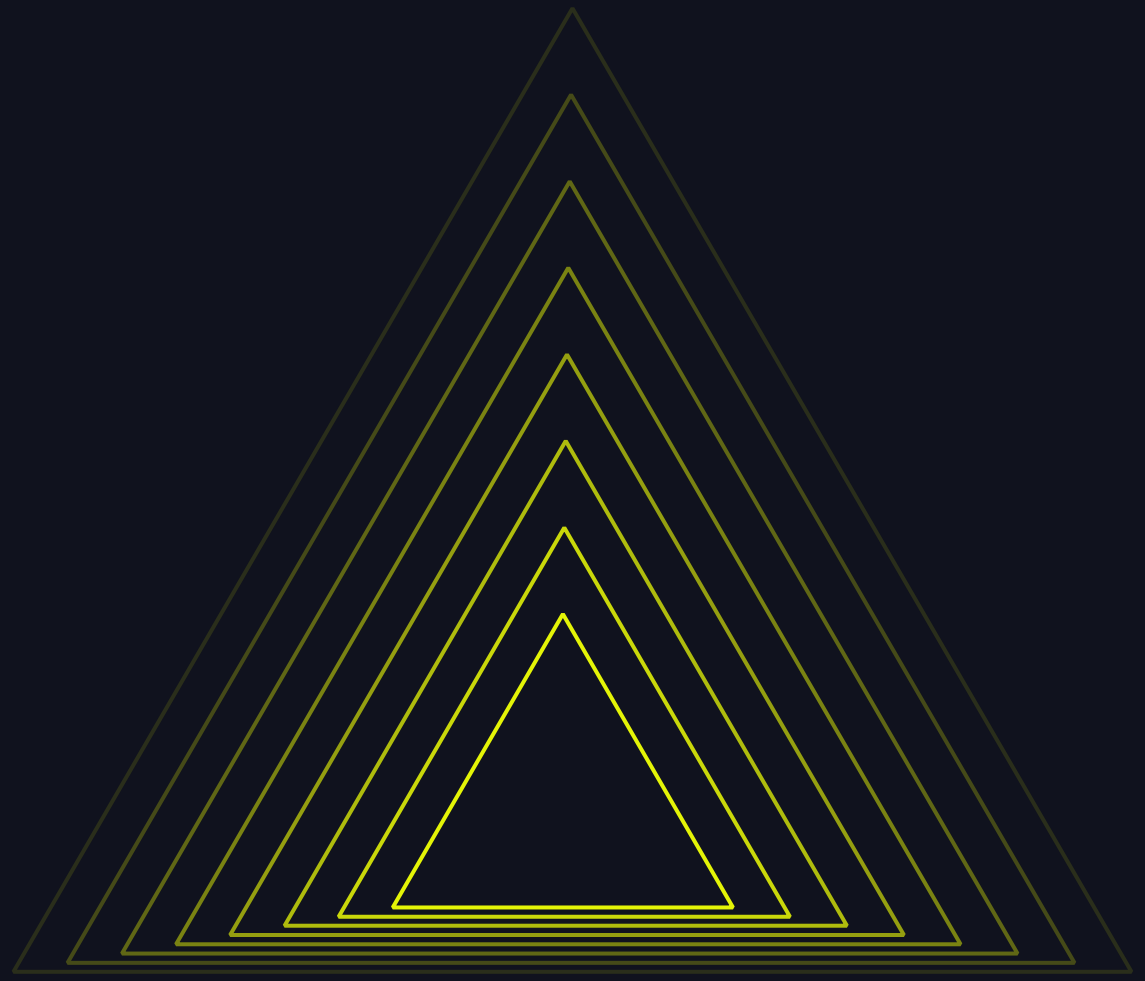
DATA+AI SUMMIT
BY databricks

Revolutionizing Data Analysis: The Shift To Databricks & AtScale



Geoffrey Roderick
Jun 2024

Founded:
1851

Headquarters:
Corning, New York

Employees:
~50,000 worldwide

2023 Revenue:
$12.6 billion

Fortune 500 Ranking (2023):
292

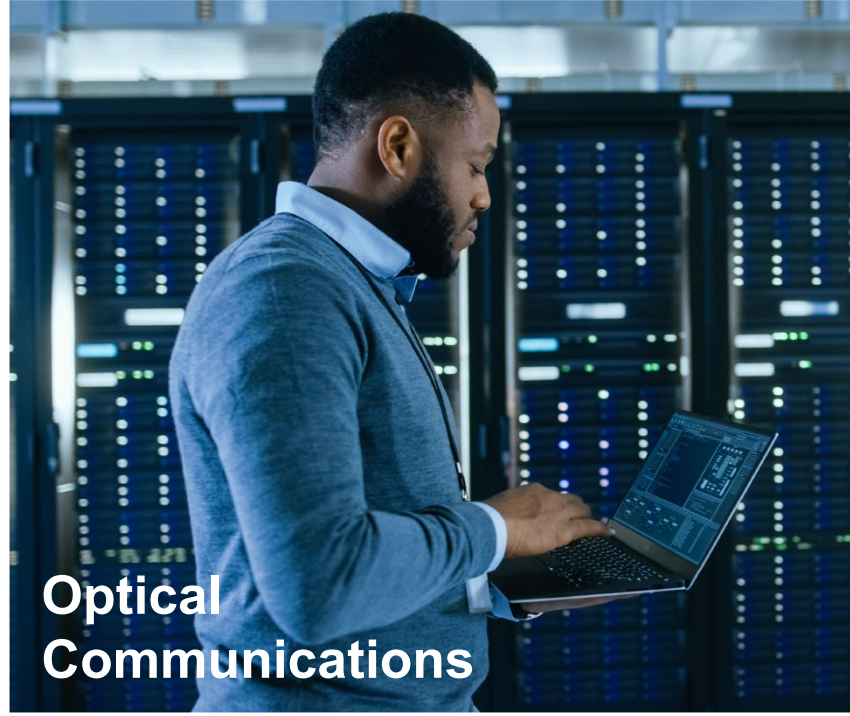
Corning Incorporated is one of the world's leading innovators in materials science. For 170 years, Corning has applied its unparalleled expertise in glass science, ceramic science, and optical physics to develop products and processes that have transformed industries and enhanced people's lives.

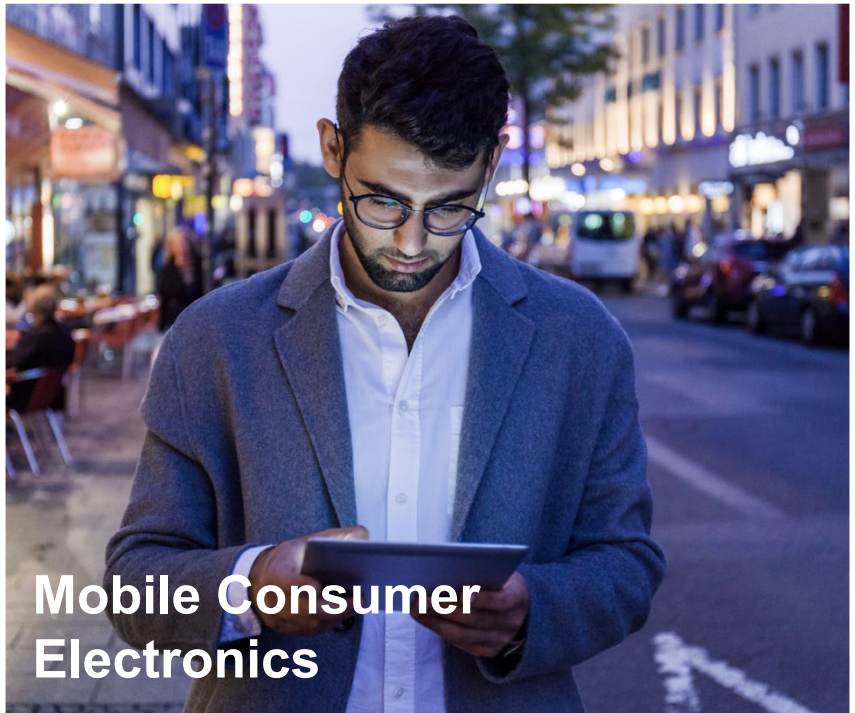
CORNING

Corning’s mission is another 170 years of innovation and independence. But we never take our eyes off the big picture. Our ultimate goal is a world that is better because of our efforts: a world with cleaner air, healthier lives, richer entertainment experiences, and more efficient communication.

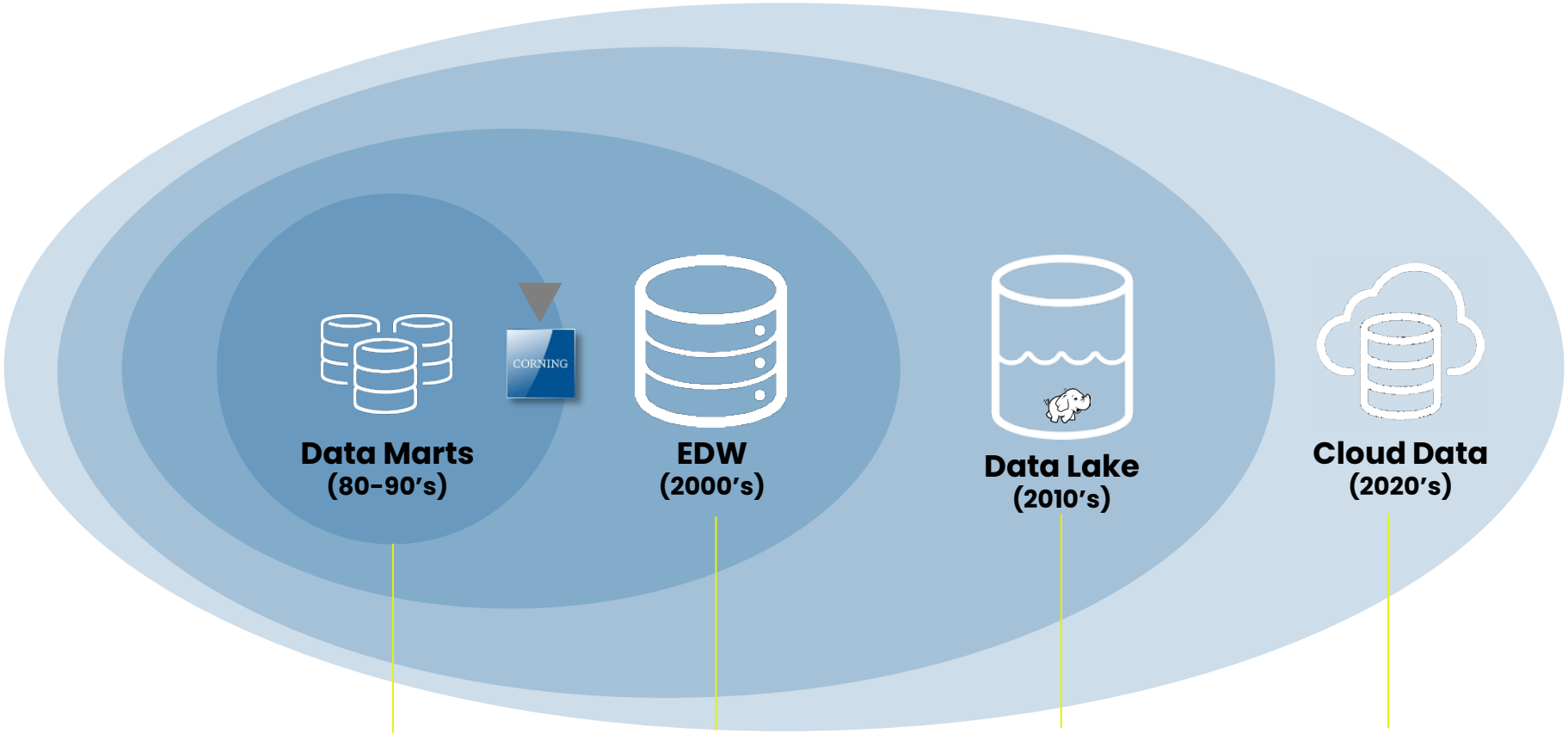
https://www.corning.com/media/worldwide/global/documents/corning_at_a_glance.pdf

The industries we help shape



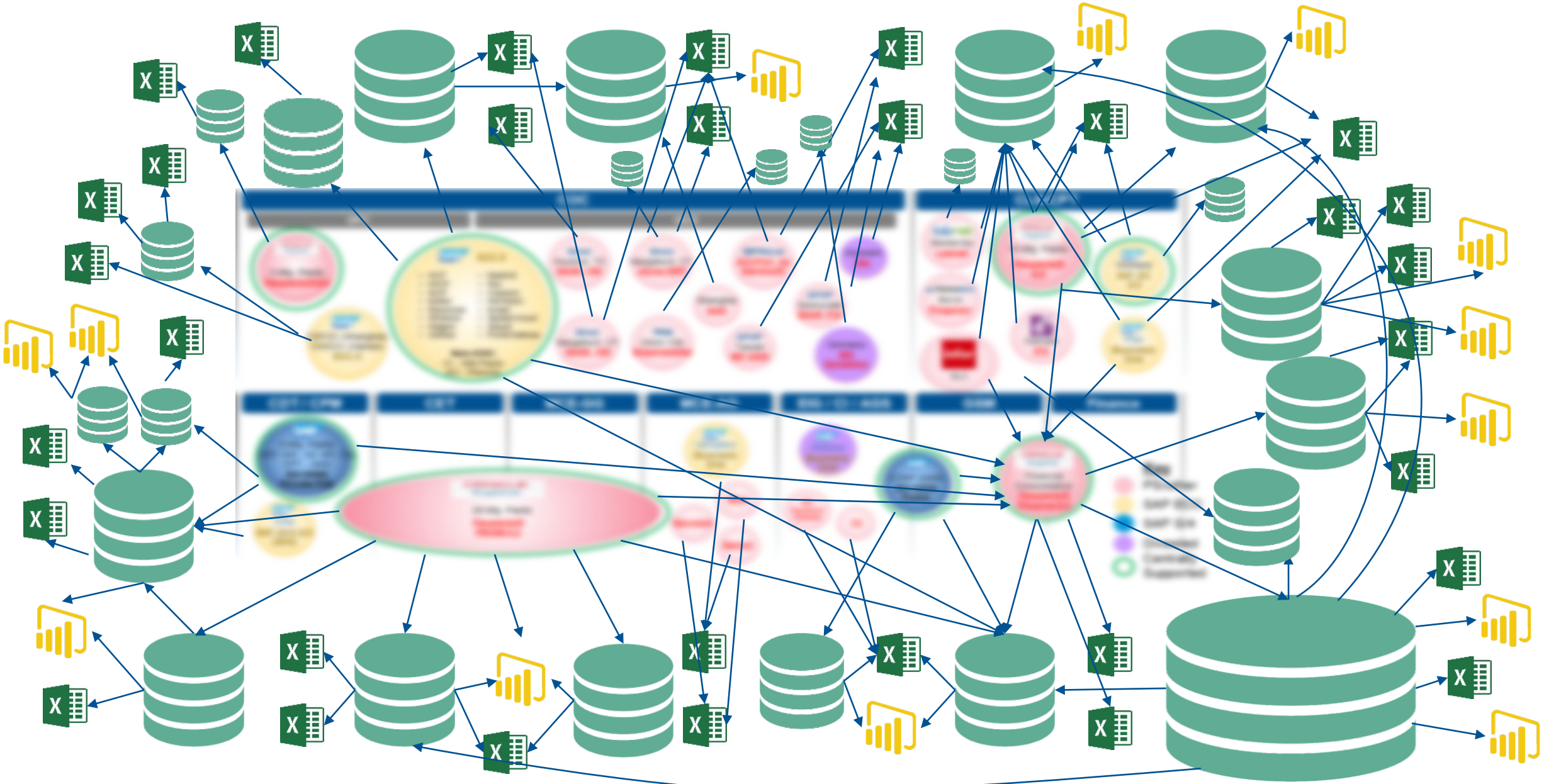


History of Data Processing



ERP Landscape

Current State



Current State

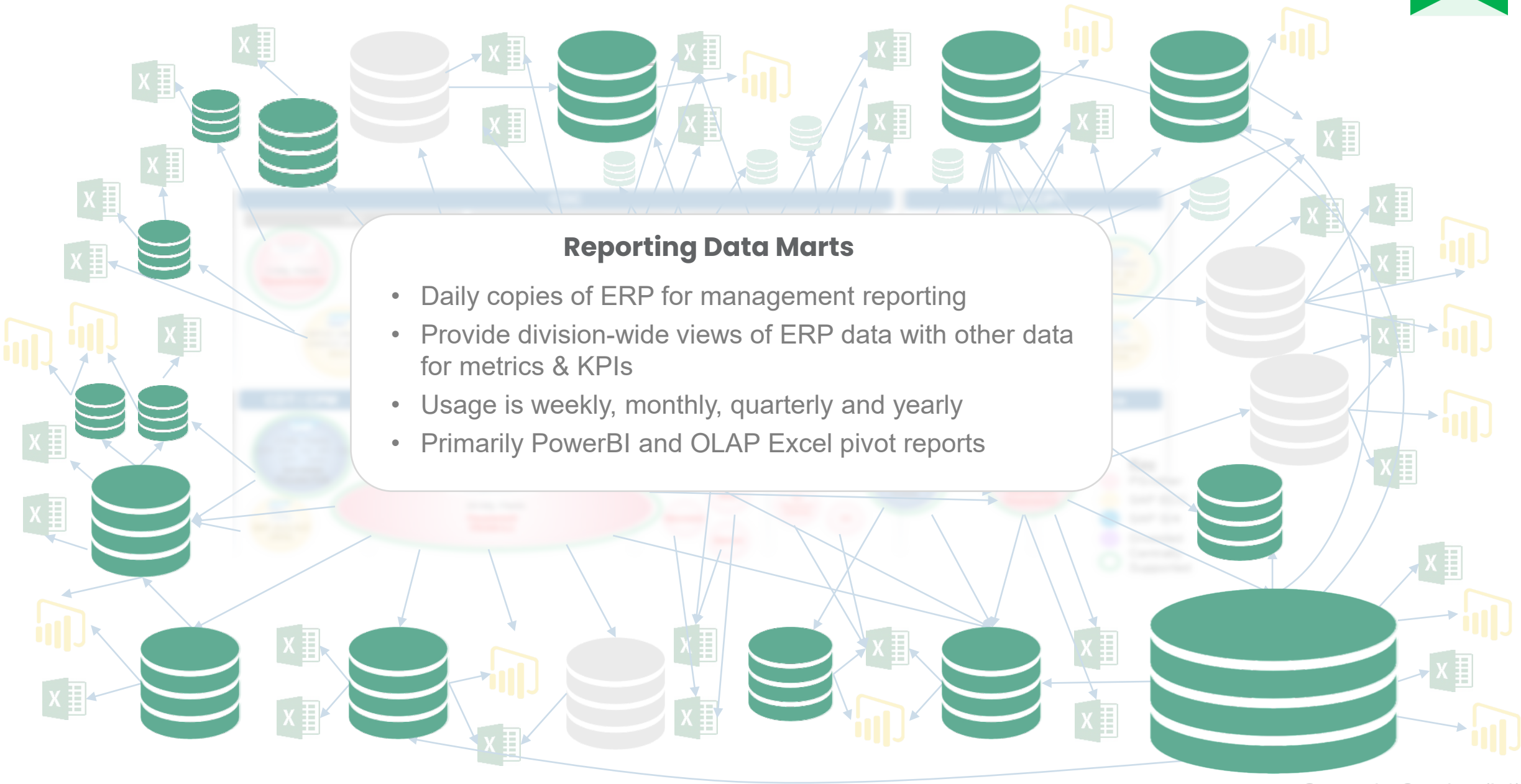
Reporting & Analytics

Operational Data Stores

- Replicated copies of the ERP for transactional reporting
- ERP could not support query volume
- Supports ad-hoc & intraday reporting
- Usage is hourly, daily weekly, monthly
- Largely Excel oriented

Current State

Reporting & Analytics

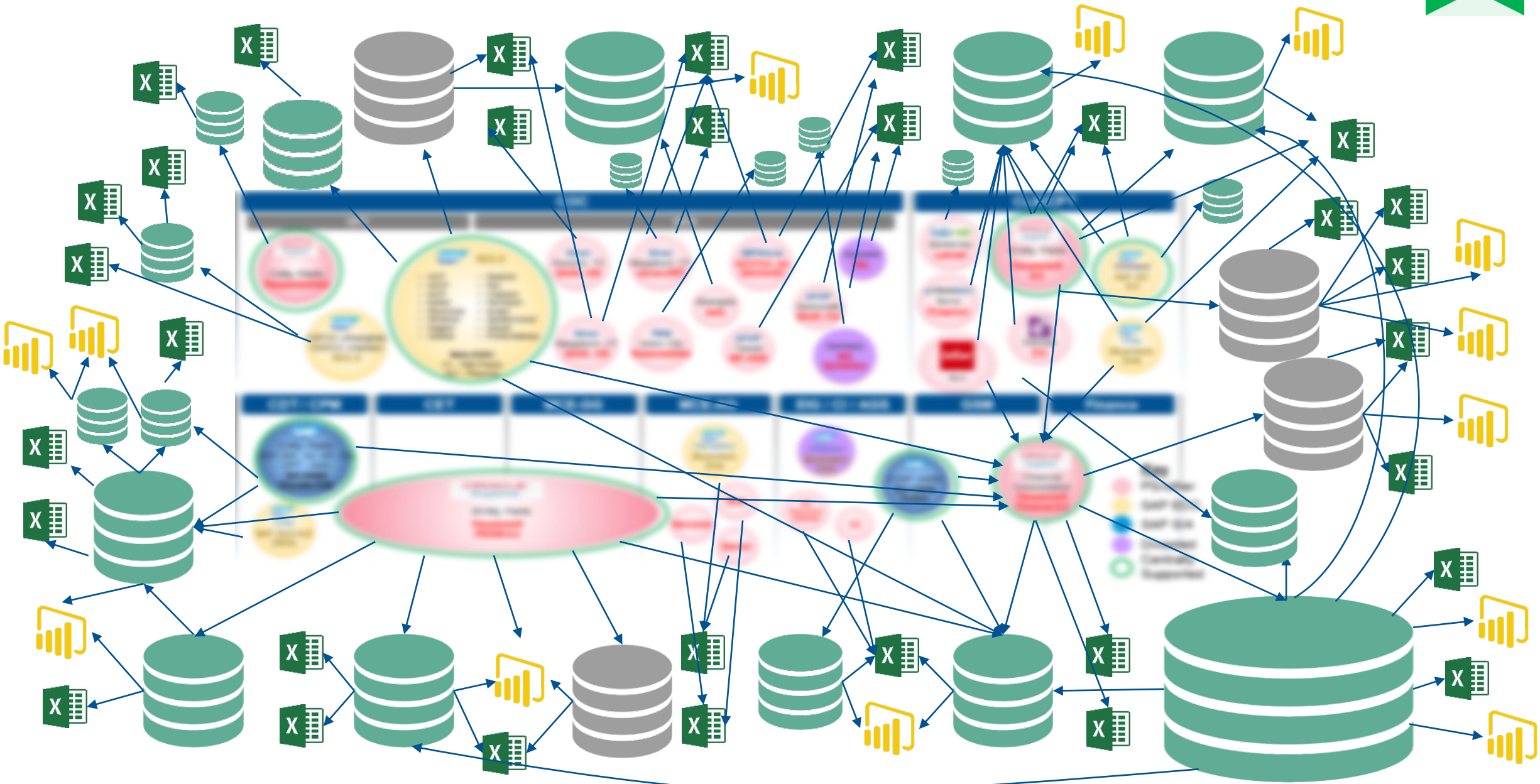


Reporting Data Marts

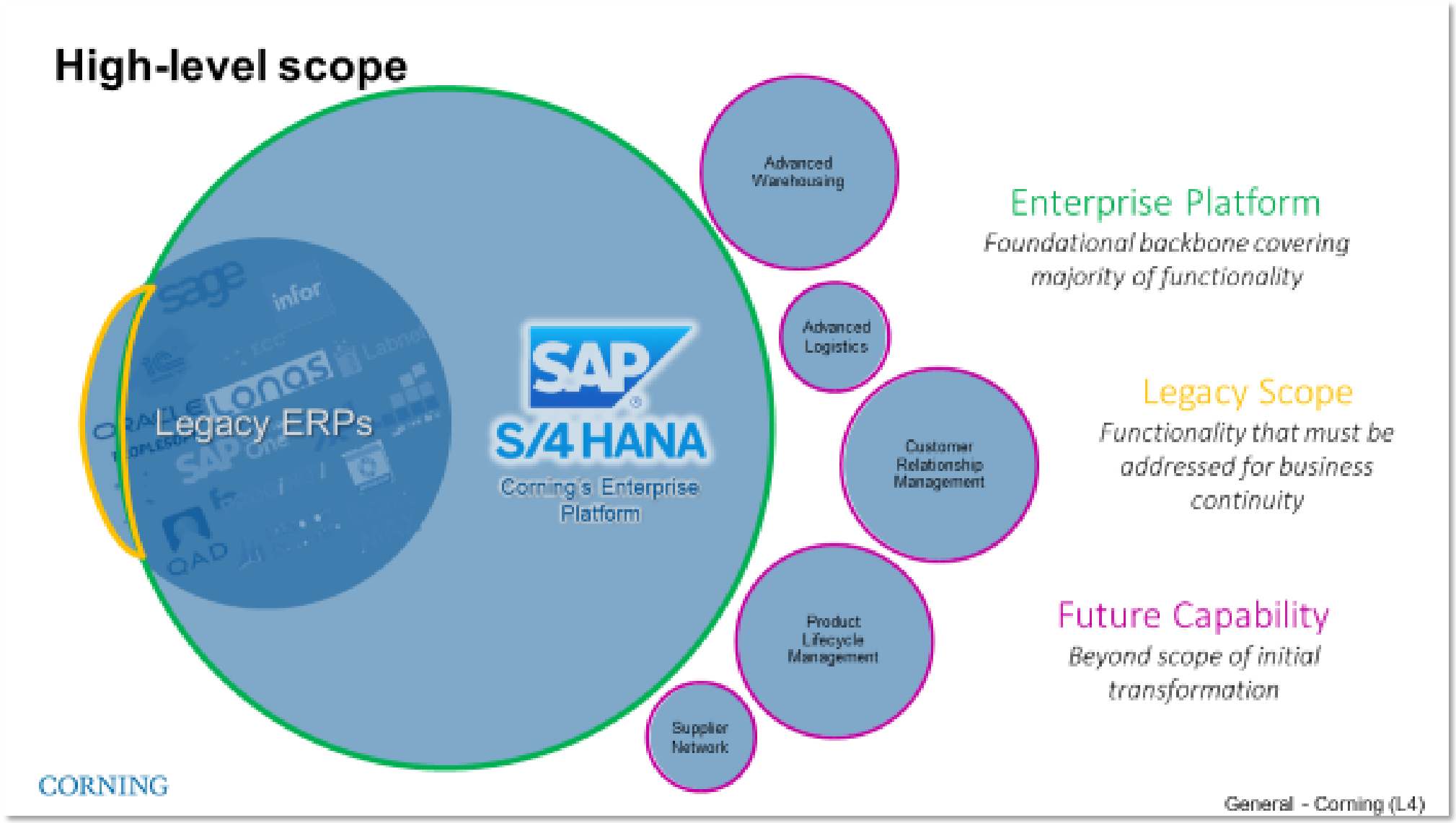
- Daily copies of ERP for management reporting
- Provide division-wide views of ERP data with other data for metrics & KPIs
- Usage is weekly, monthly, quarterly and yearly
- Primarily PowerBI and OLAP Excel pivot reports

Current State

Reporting & Analytics

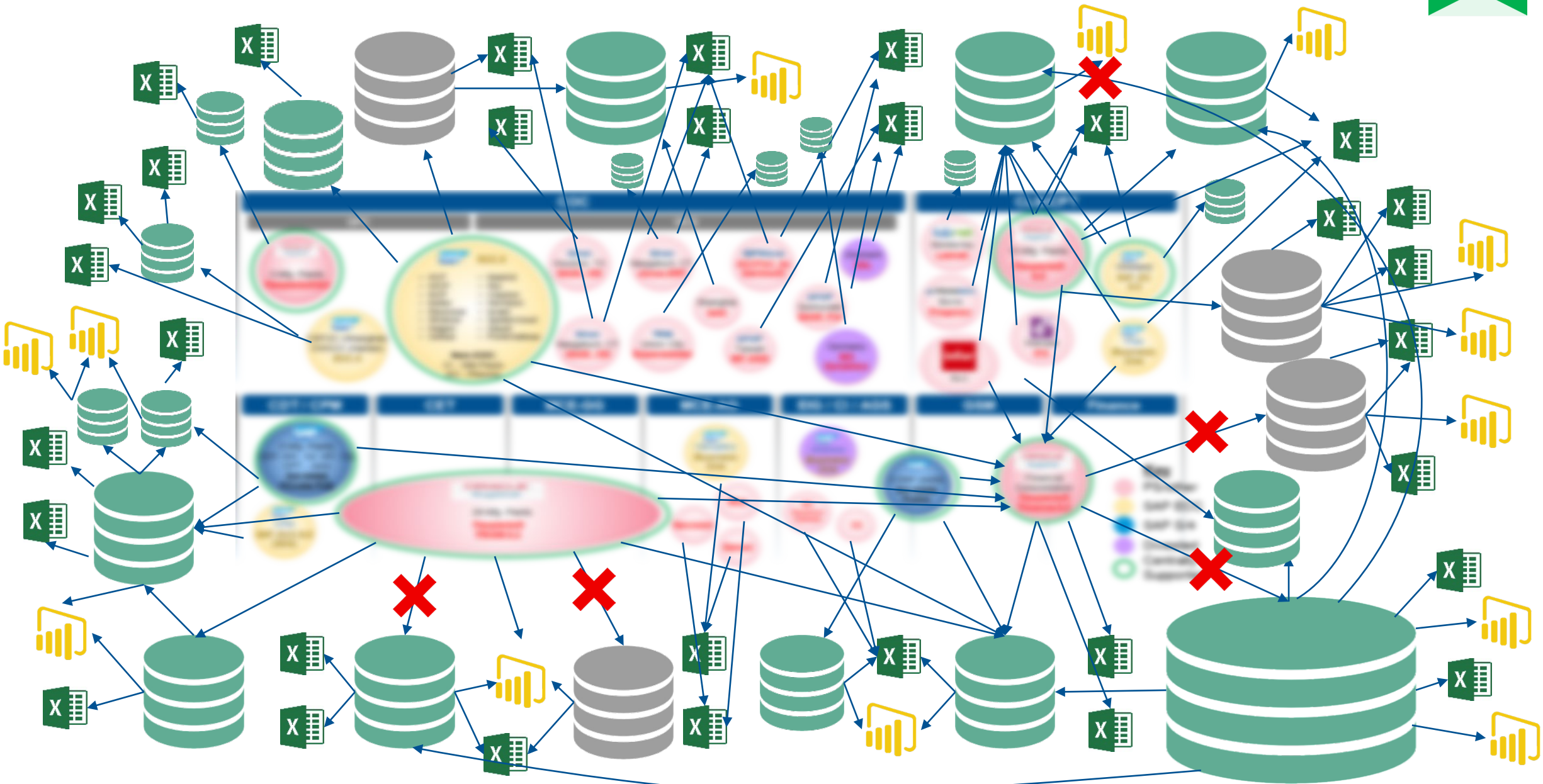


ERP Transformation



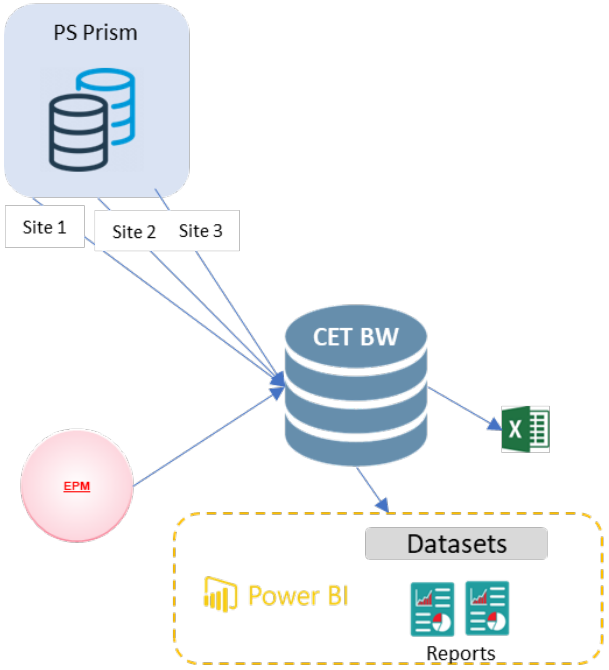
Current State

Reporting & Analytics

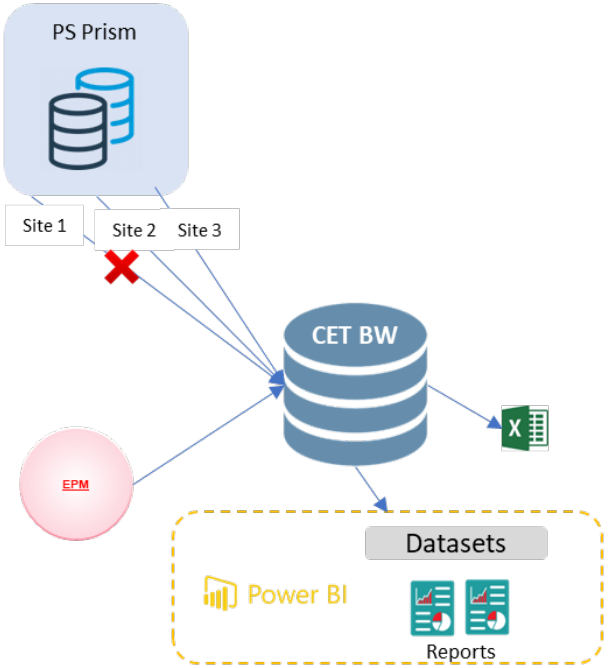


Our Current Data Marts Will Break With ERP Transformation

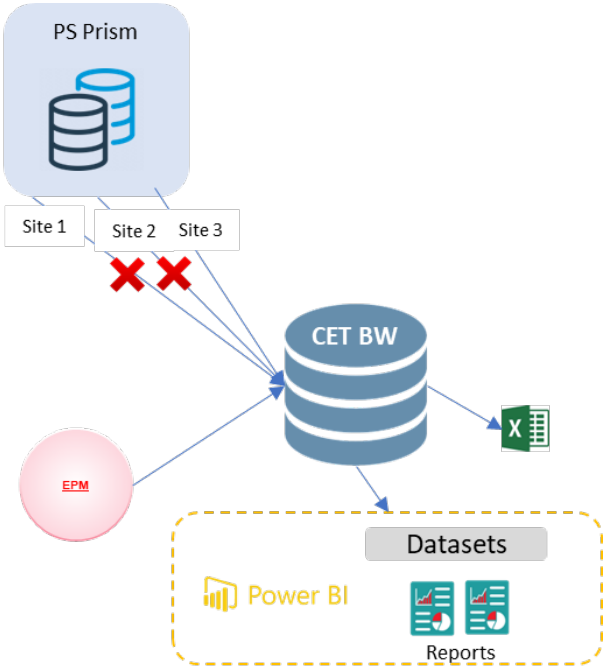
Division Example



1st SAP Deployment



2nd SAP Deployment



Options

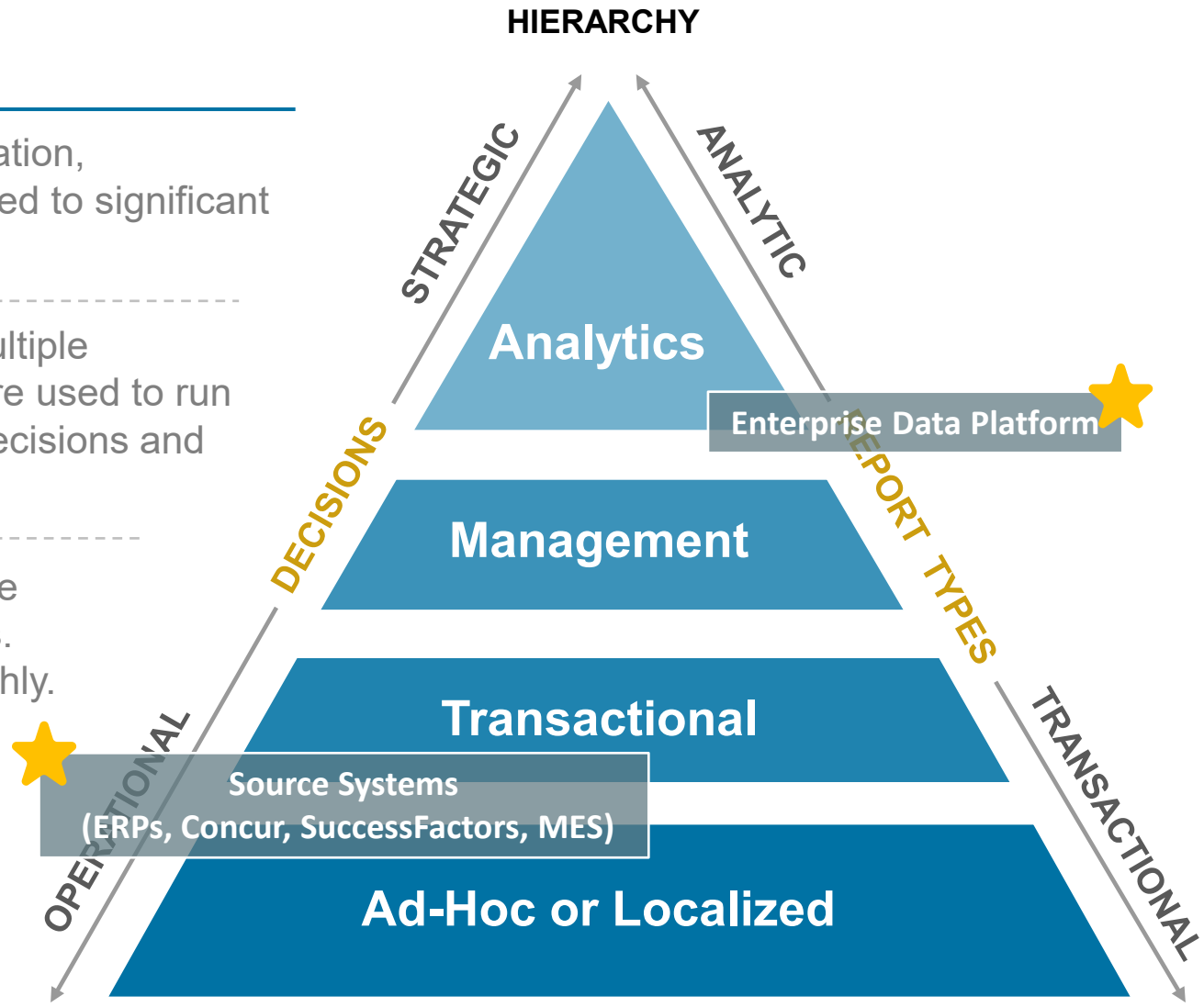
Option	Pros	Cons
1 **Do Nothing**	Least expensive approach Leverage reporting out of SAP	Lose division-wide reporting during transition No historical data from legacy ERPs in SAP Significant manual effort to compile reports
2 **Backfeed**	Preserves existing reporting during transition	Time consuming and expensive Some data won't fit backwards Each division would have to do the same exercise High maintenance costs
3 **Enterprise Data Platform** 	Aligns with overall data strategy Common approach for all divisions Integrate with data from other systems	Requires strategic investment Effort to provide division-wide reporting during transition Significant change management May require defining standard ERP report catalog

Guiding Principles

1. **Modern data bases architectures support OLTP and reporting** on single instance
2. **Do reporting in the first system** that has all the data needed
3. **Leverage inherent reporting capabilities** where possible
4. **Only move data for reporting** when a system cannot support the requirement or to add value (history, enrichment, etc.) to the data
5. **Don’t migrate non-operational, historical data** to new transactional systems just for reporting
6. **Stitching data together across different systems** should be done in a centralized Enterprise Data Platform (EDP).
7. **Analytics with intensive business rules or algorithms** should be done in the EDP.
8. **Business rules should be done at the data layer and not reports** for reusability
9. **Bring the user and processing to the data,** not the data to the user or processing
10. **Let natural disruption drive re-platforming,** not re-platforming for re-platforming sake

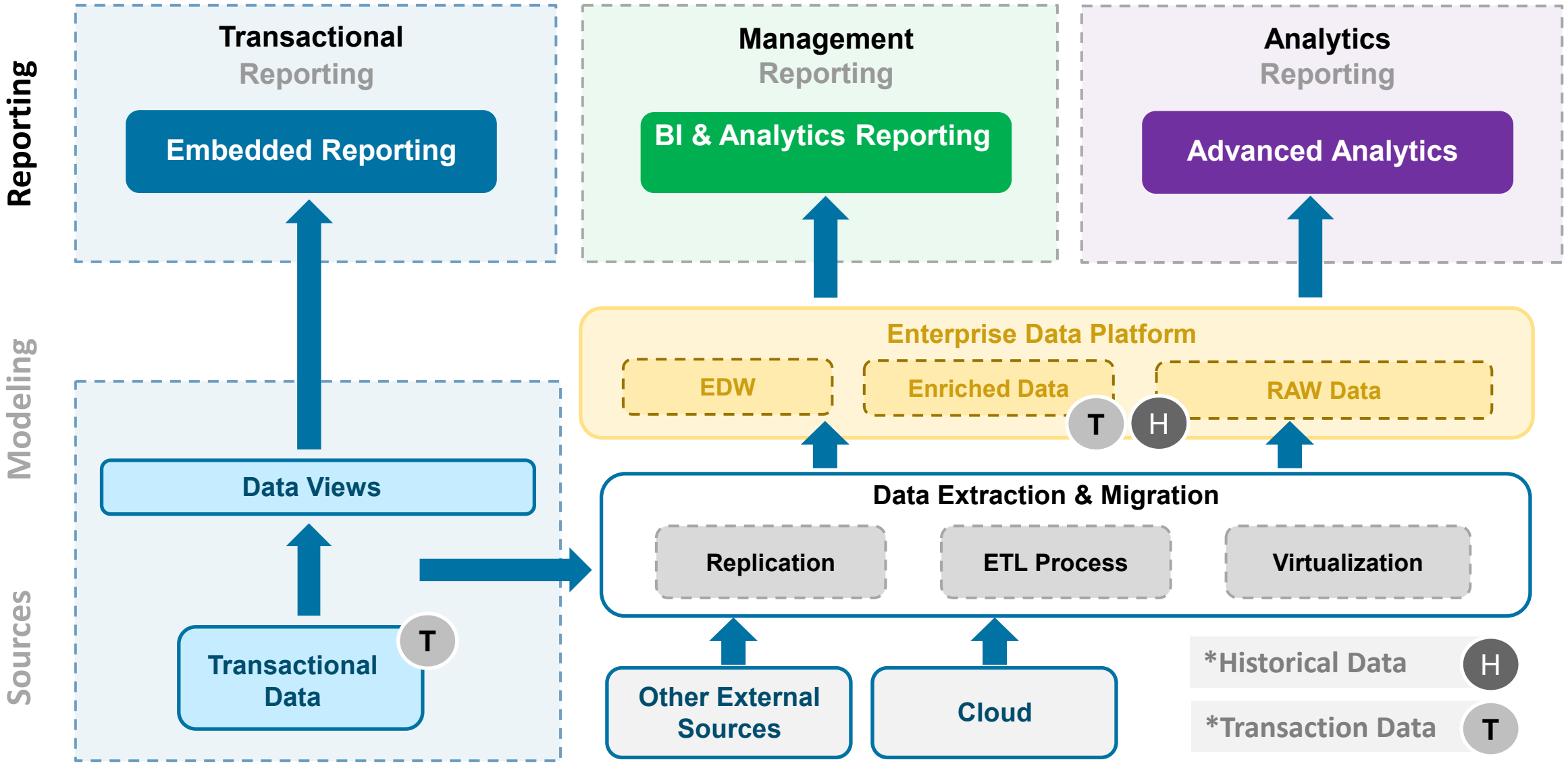
Data & Analytics: Reporting Type Hierarchy & Categories

REPORT CATEGORY User Community	INFORMATION REQUIRED
Analytics Analytics Users	Supporting the discovery, interpretation, communication and decisions related to significant patterns and insights from data.
Management Executives and Senior Managers	Reports that combine data from multiple systems that generate KPI's that are used to run the organization, make business decisions and to monitor progress.
Transactional Business Users	Intraday reporting that highlights the details of a company's transactions. Usage is hourly, daily weekly, monthly.
Ad-Hoc Technical Users	Ad-hoc or local reports that display data for personal analysis or department metrics.

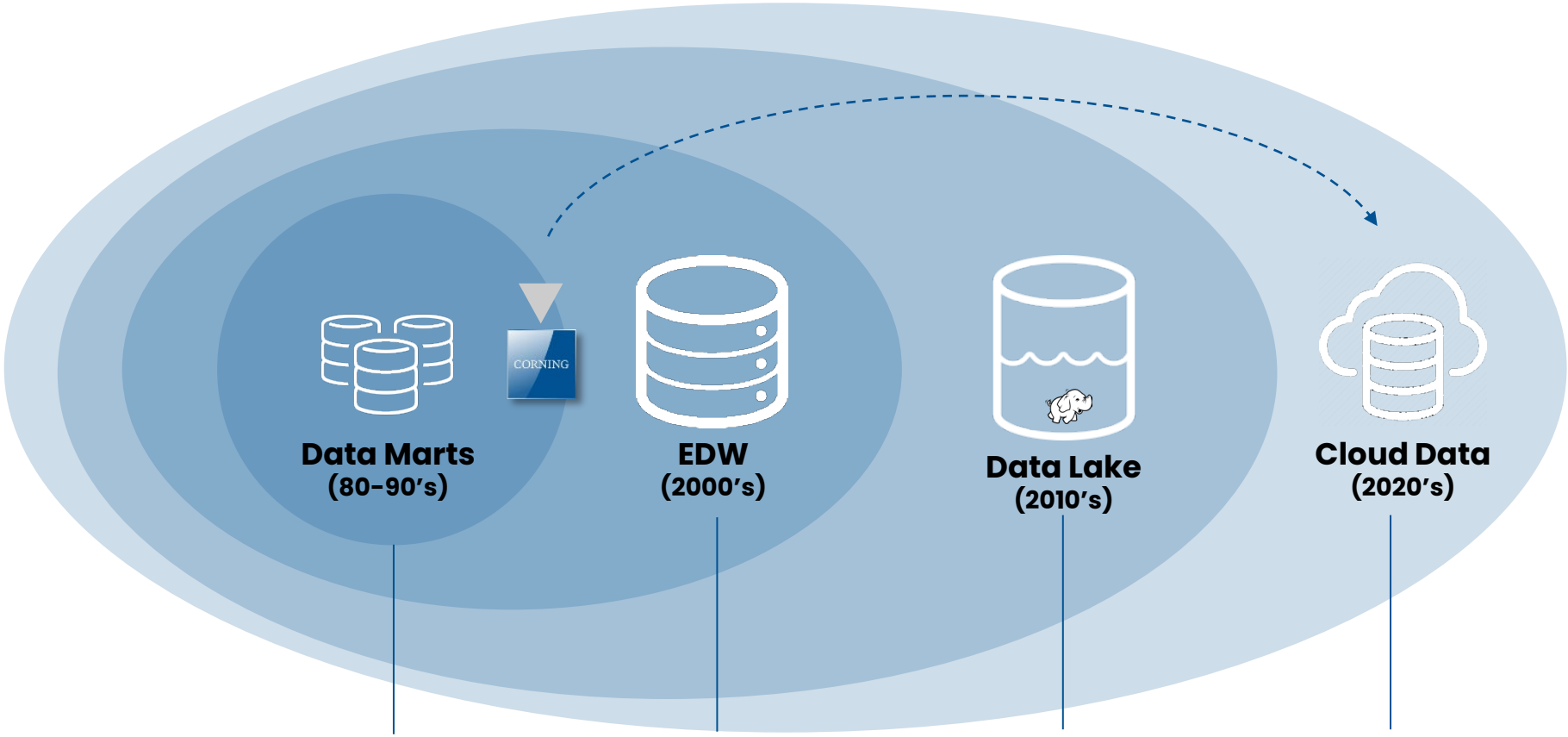


Data & Analytics Reporting: Defining the Boundaries

Reporting & Analytics

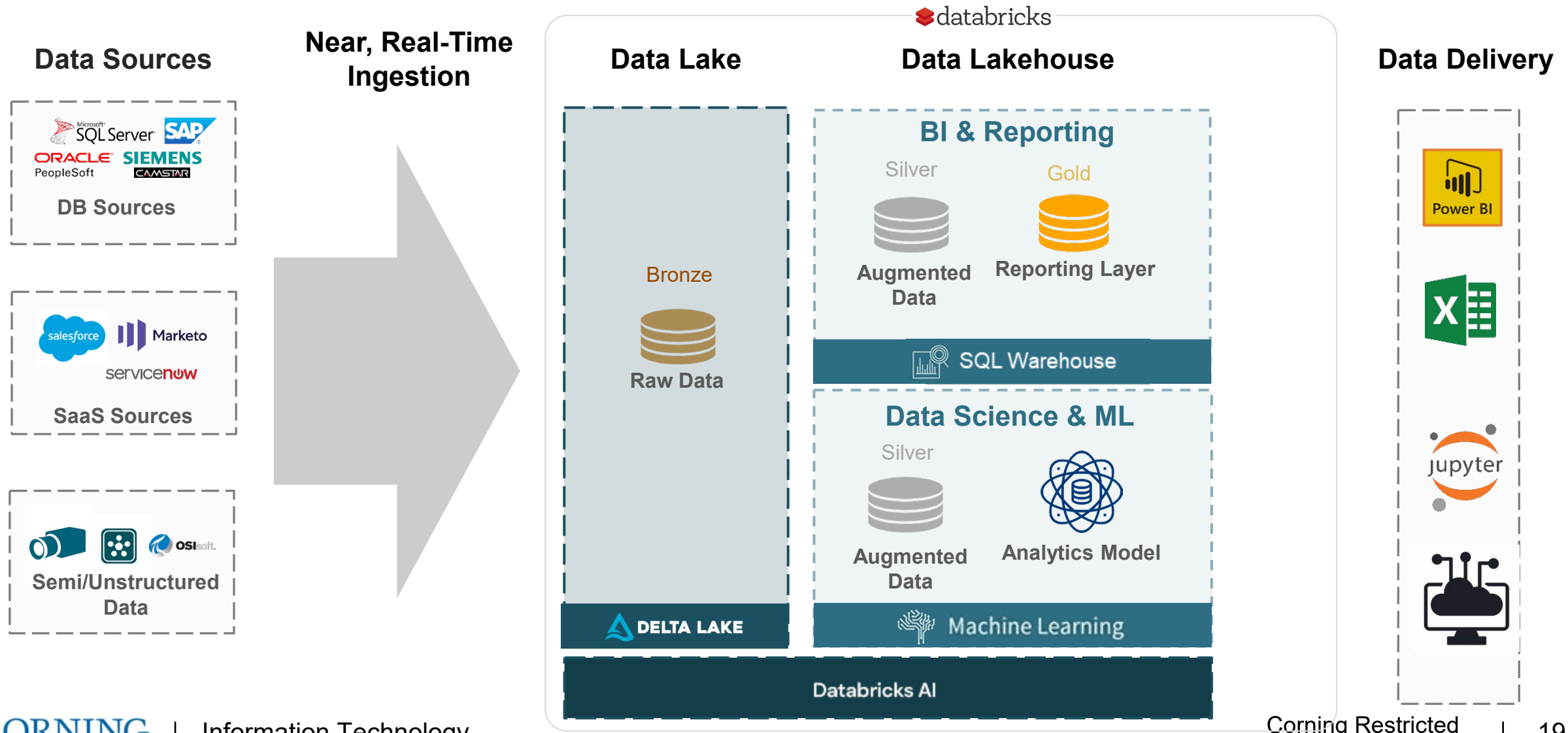


History of Data Processing

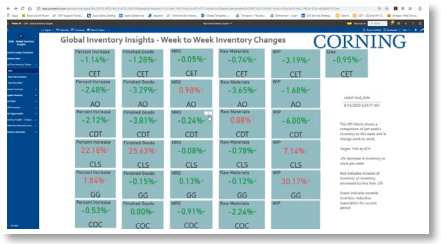


Enterprise Data Platform Architecture: Data Processing Framework

This framework adheres to a land once data architecture principle

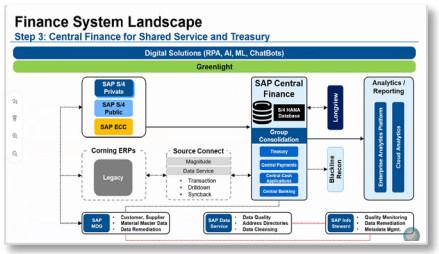


Embedding Business Strategies With Data Transformation



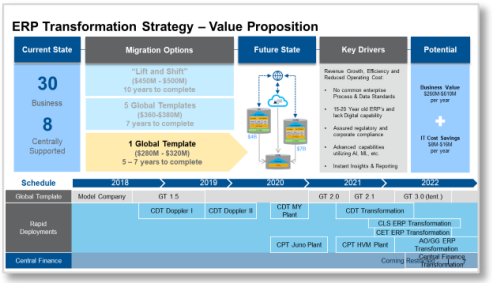
Inventory Insights

Rapid data acquisition from SAP on-prem & cloud sources



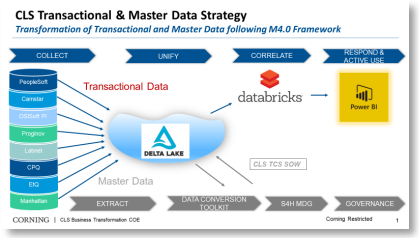
Finance Transformation

Accelerate ingestion from SAP, PS & other sources



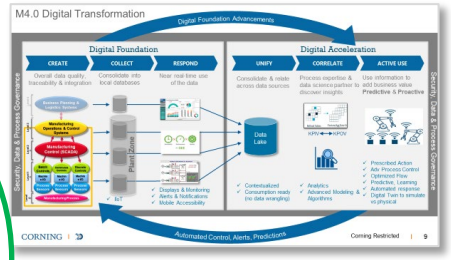
ERP Transformation

Enablement of master data and governance



CLS Data Lake

Transformation of reporting & analytics capabilities

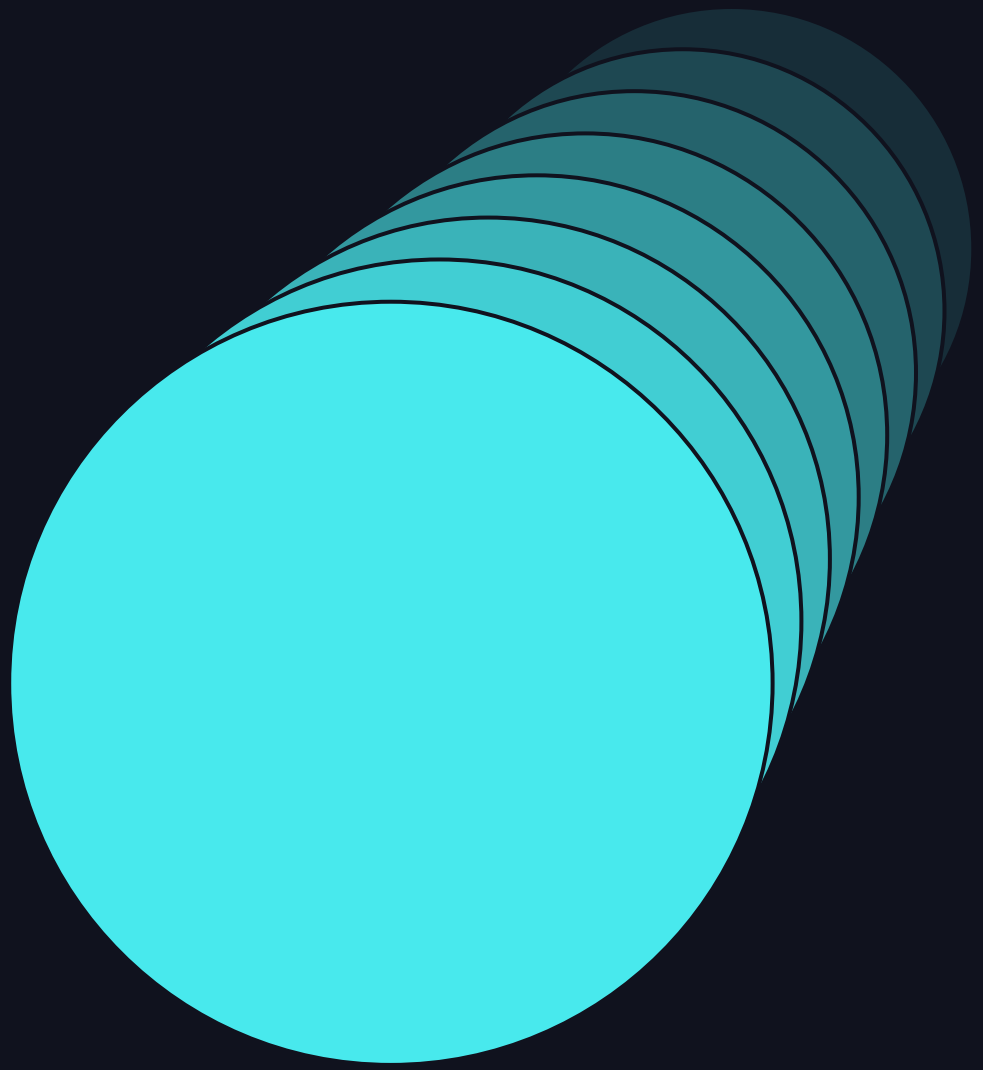


Manufacturing 4.0

Implementation of M4.0 best practices in plants & PoVs

DATA+AI SUMMIT BY databricks

Corning Life Sciences



Business Transformation COE

Enabling growth by optimizing and supporting processes, people & data

Future Landscape

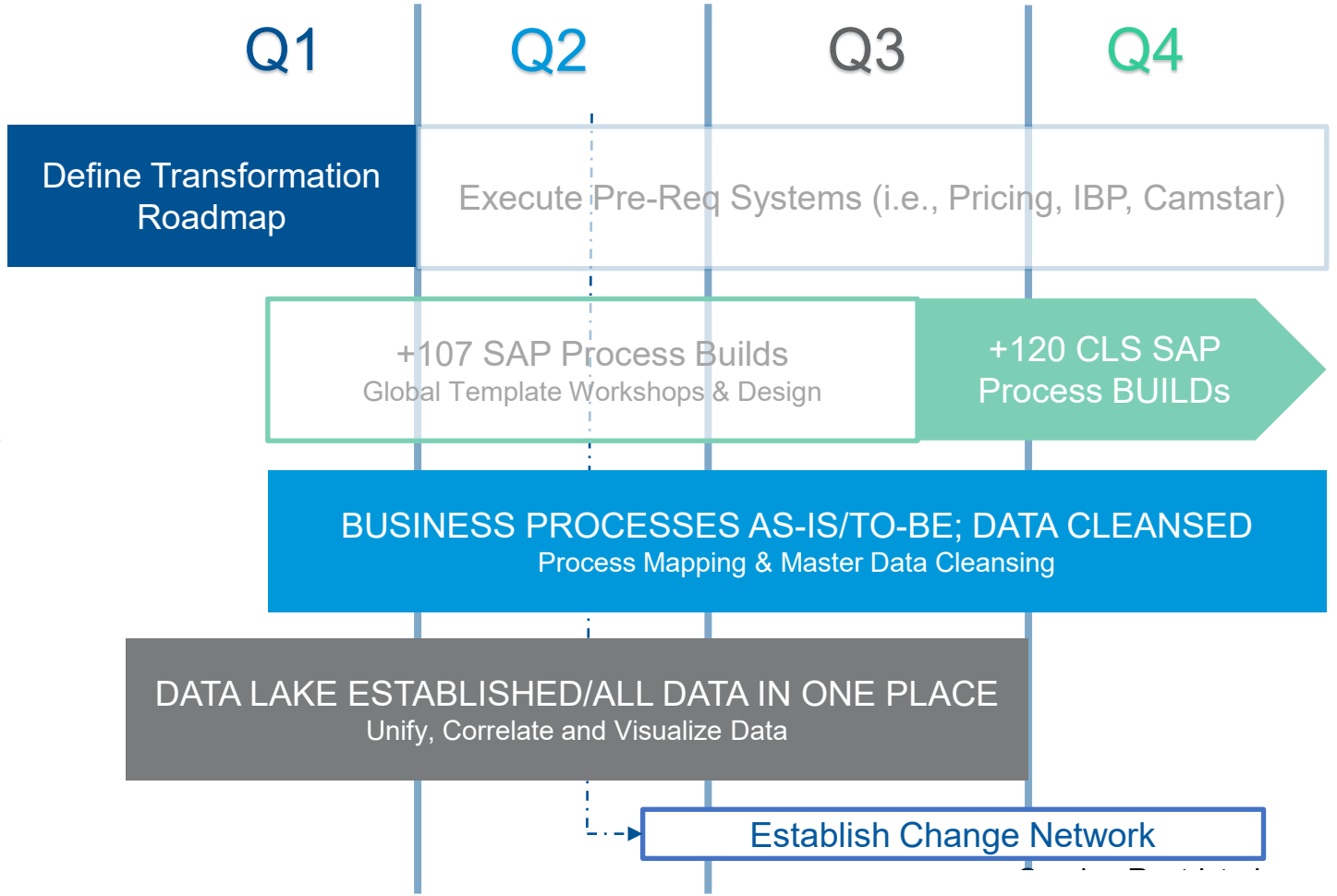
A Future Focused on the Customer

- Meet customer demands faster & achieve competitive advantage
- Employees confident in products and services they deliver to customers
- Drive consistency & scalability through governance & organizational alignment
- Quickly respond to market challenges & external forces

Business Process Optimization
Master Data & Data Governance
Business Intelligence & Data Science
Change Management

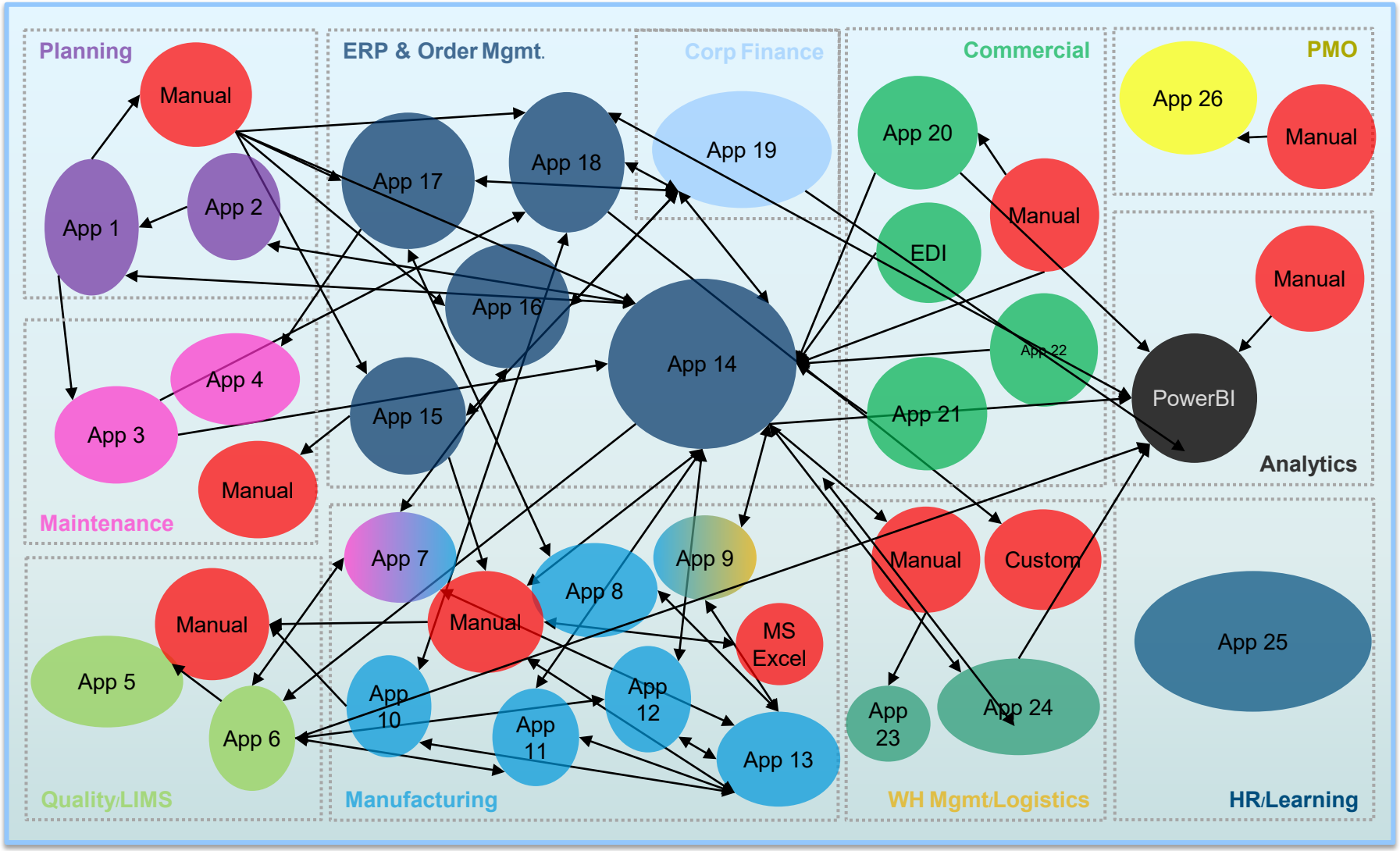
2021 Key Focus Areas

Transformation COE

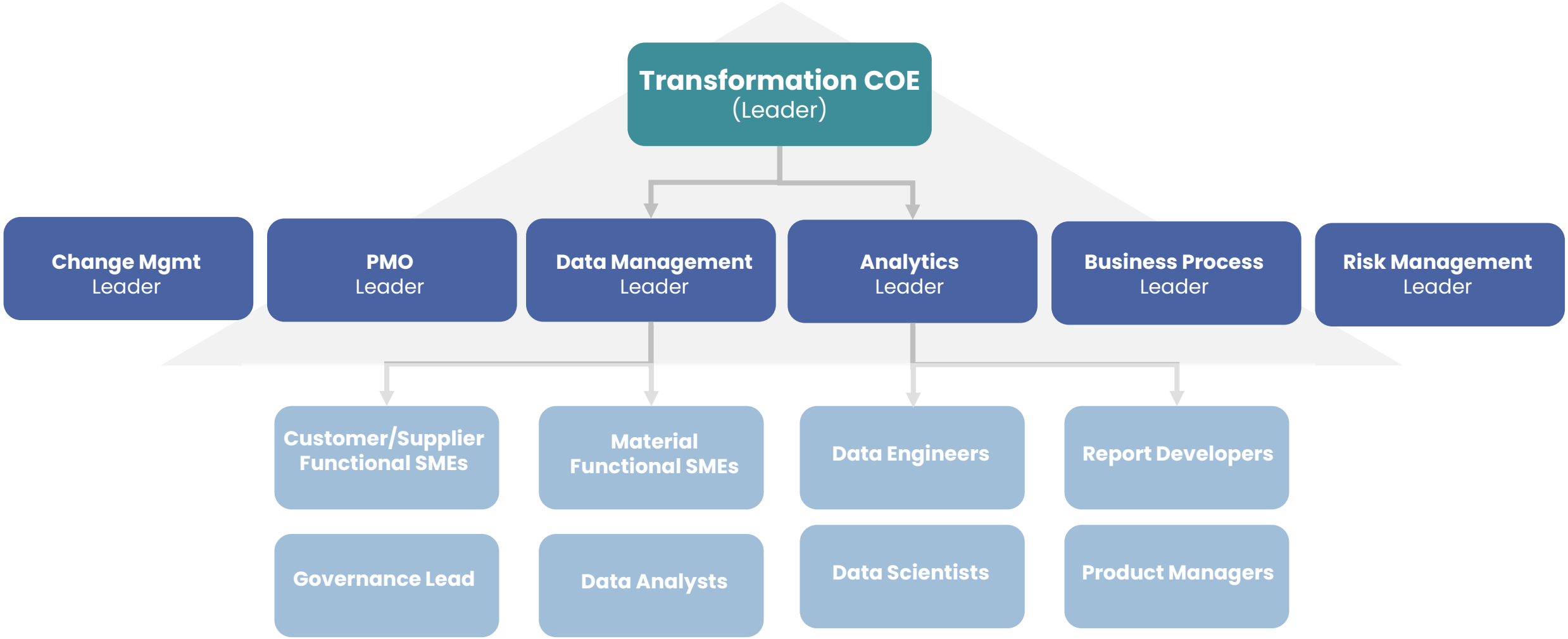


CLS Current Systems Landscape

CLS systems are disconnected, disparate, and obsolete

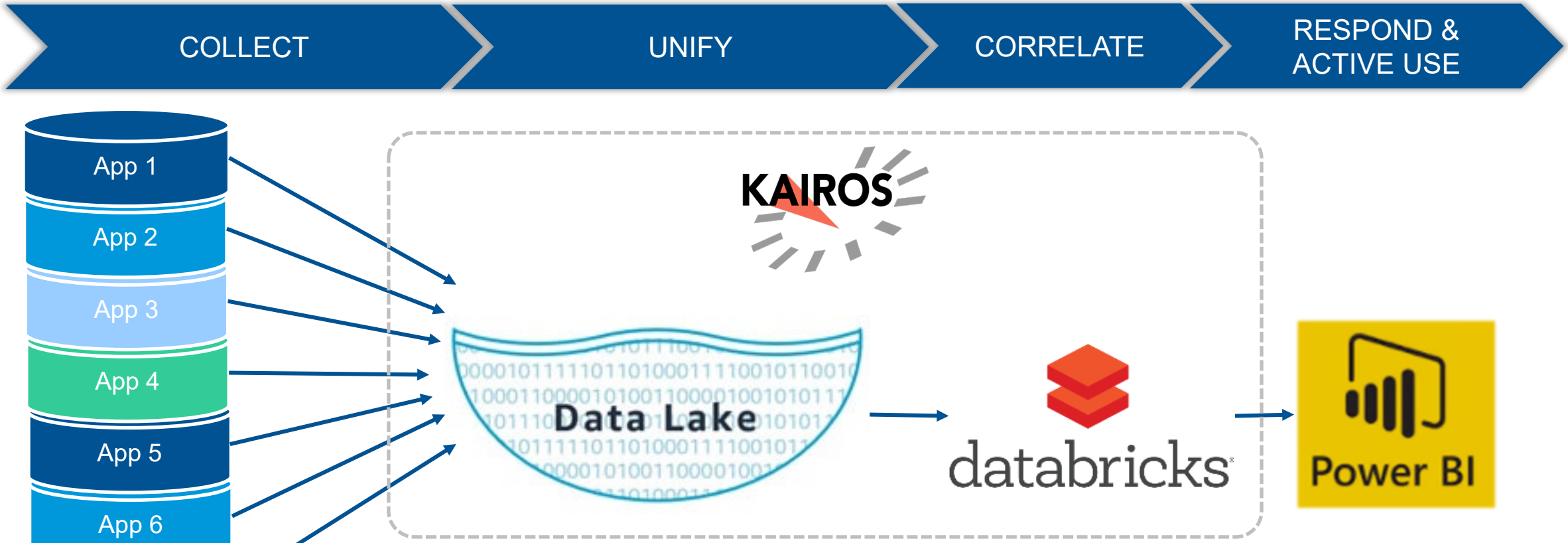


Transformation Organization – Overall Structure



CLS Transactional & Master Data Strategy

Transformation of Transactional and Master Data following M4.0 Framework



"Kairos (Greek for "right time," "season" or "opportunity") • Refers to the "timeliness" of an argument. • Often, for an ad or an argument to be successful, it needs appropriate tone and structure and come at the right time."

Single Source Truth

Business Intelligence Platform centralizes how Life Sciences accesses data

Intake

Metrics & Insights

Power BI

SharePoint

Welcome to the CLS BI site!

This site is packed with business intelligence and business insights resources for CLS worldwide.

Do you know who we are? Click Here to find out more about us!

- Connected Customer
- Customer Fulfillment
- Financial Reconciliation
- Innovation & Execution
- People Support

Training

Using PowerBI
Understanding Data Models
BI Platform Navigation
Data/Report Interpretation
and more…

Communication

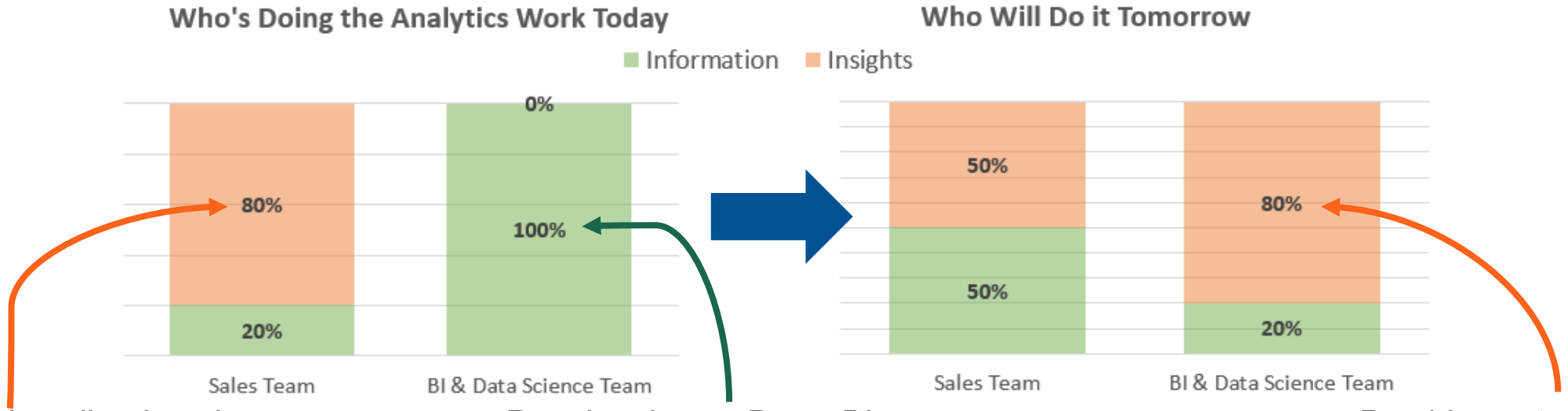
Centralized user experience designed around core operations

Standardized external & internal performance understanding

Foundation for becoming more *proactive* using our data

Information vs. Insights

Where we spend our time today vs. where we need to spend our time



Performing siloed analyses, taking time away from their core responsibility

Requires heavy PowerBI work to slice & dice data 100s of ways

Provide centralized tools to surface insights that are broadly beneficial

This change is aligned with capabilities we hire for…

- Meet or exceed your budget
- Build relationships with customers
- Build, implement and maintain a sales plan
- Know your products, market, channel, etc.

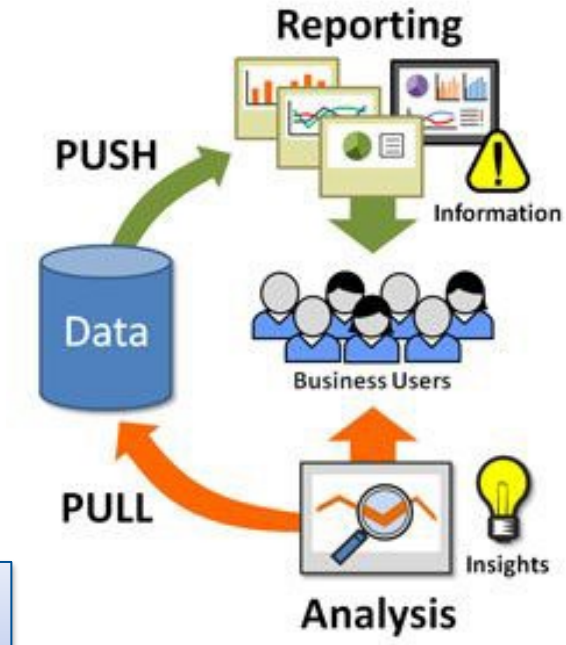
- Design, develop, & maintain analytics solutions to improve business performance
- Prepare and deliver analyses and insights

Information vs. Insights

Information is only actionable when you can find insights

Information Access

- Add a filter for new products
- Let me export the raw data
- Add product class to the hierarchy



Business Needs

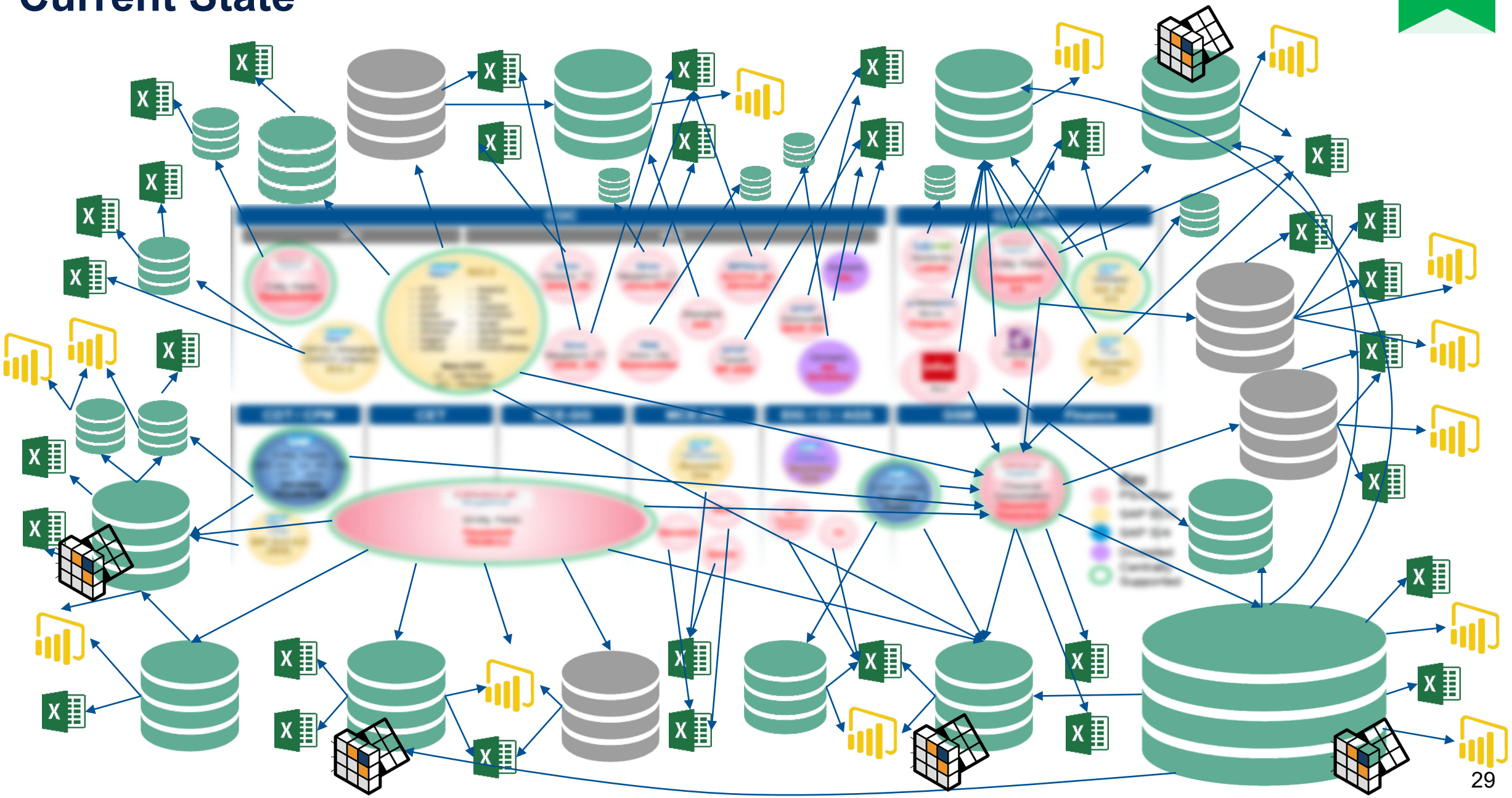
- How can I quickly identify accounts at risk?
- Need to easily identify ways to increase revenue through cross-sell opportunities

Individual users performing siloed analyses, takes time away from their core responsibility

Centralized analytics tools to surface insights that are broadly beneficial

Current State

Reporting & Analytics



Maintaining the OLAP Data Experience

Problem Statement

- End-user feedback indicated that slice and dice through Excel was a critical data analysis capability
- Adoption of Databricks would be hindered without providing parity to with existing data marts
- Excel can connect to Databricks but doesn't support seamless multi-dimensional analysis

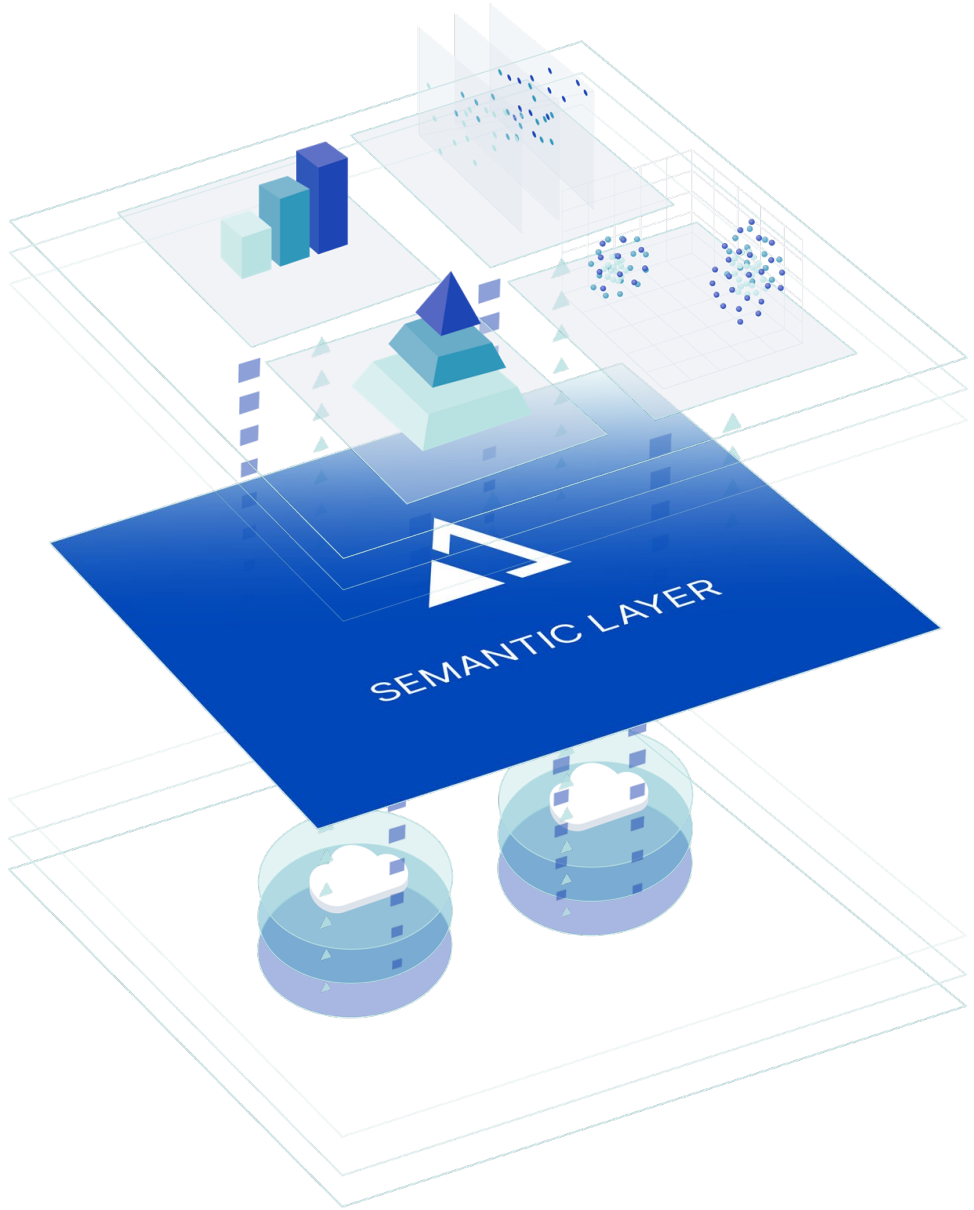
Solution Exploration

- Databricks SQL Warehouse – didn't fully meet the need
- Snowflake – data latency, integration, and capability limitations
- AtScale - runs natively inside of Databricks and provides the ability to build multi-dimensional views

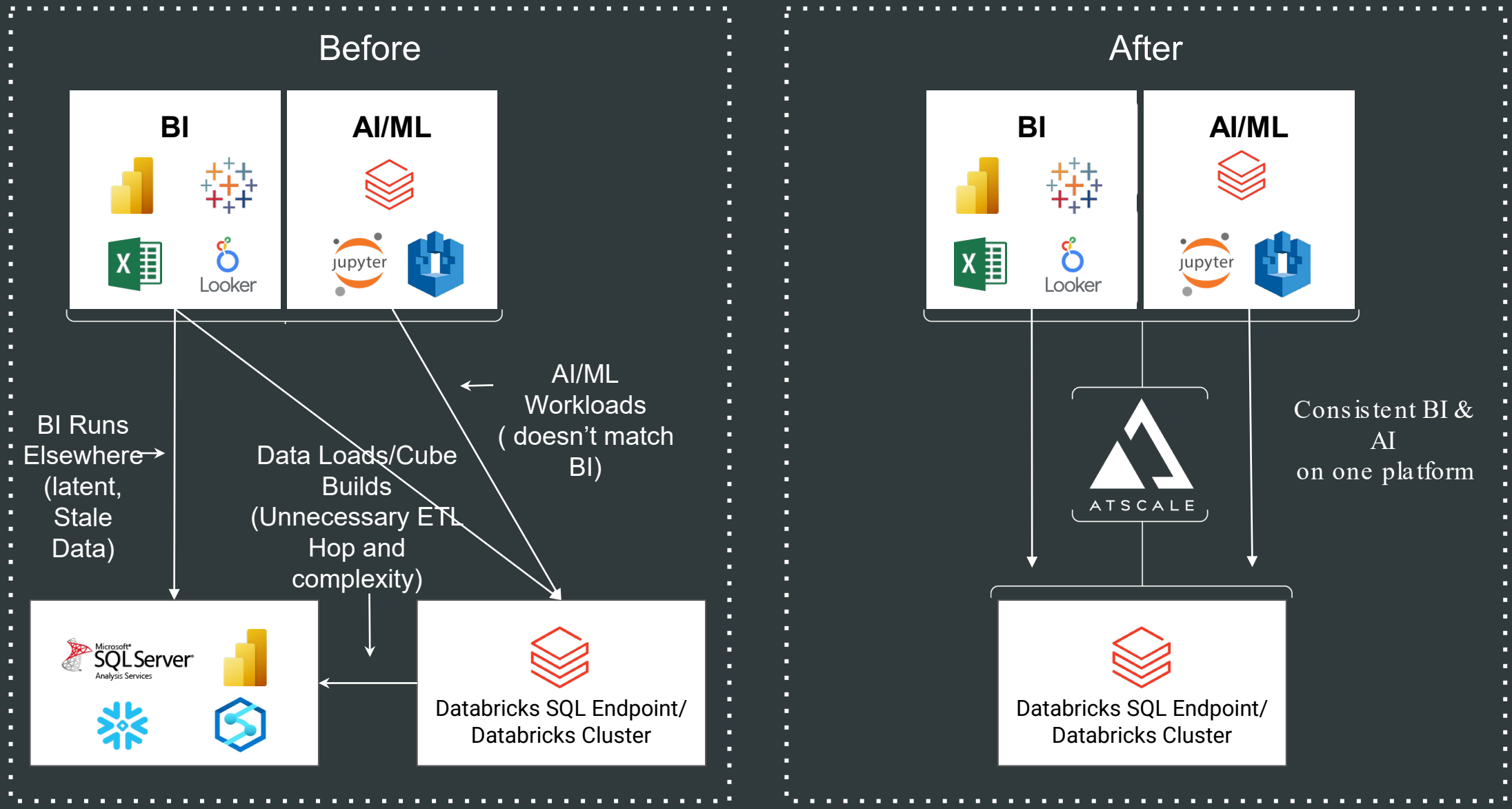
AtScale Vision

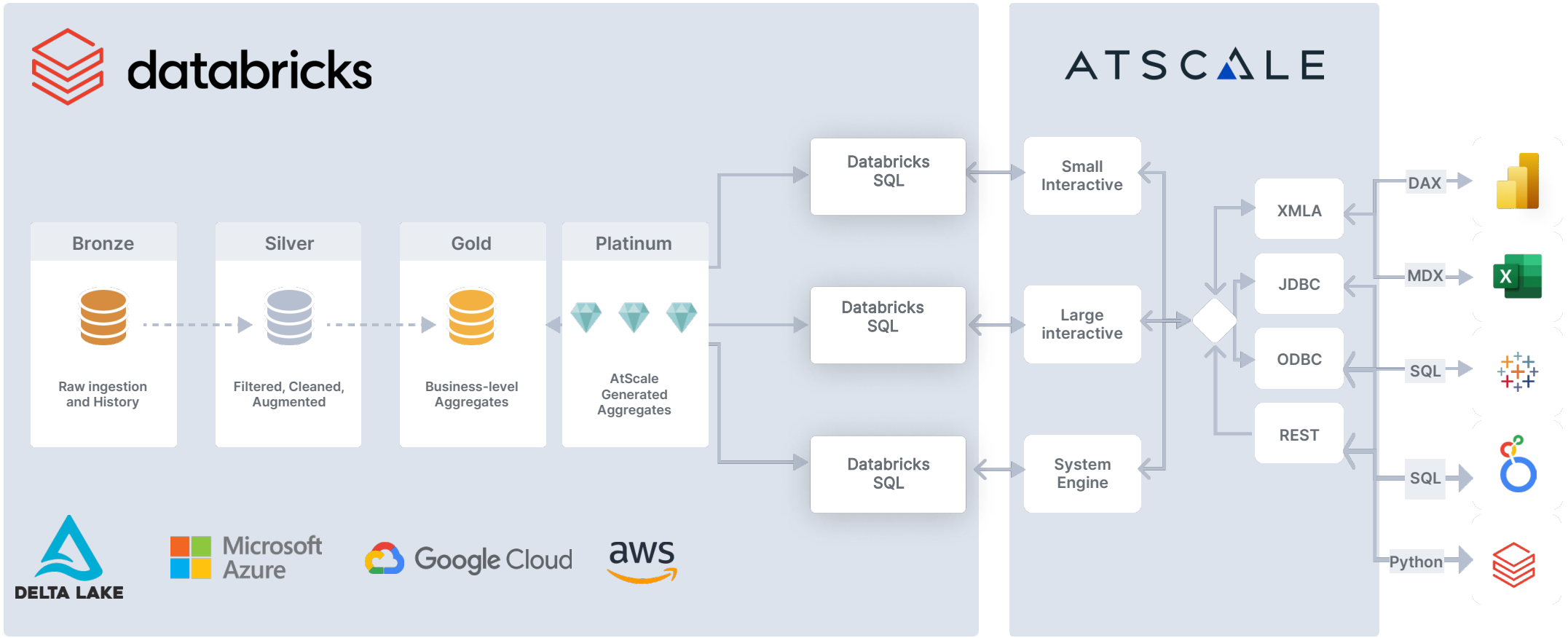
Deliver the leading universal semantic layer platform for enterprise data teams to manage and scale analytics infrastructure for Business Intelligence, Generative AI, and Data Applications.

- AtScale does not move data off cloud data platforms.
- AtScale leverages existing BI and data infrastructure.
- AtScale integrates openly with data fabric for managing analytics metadata and governance.



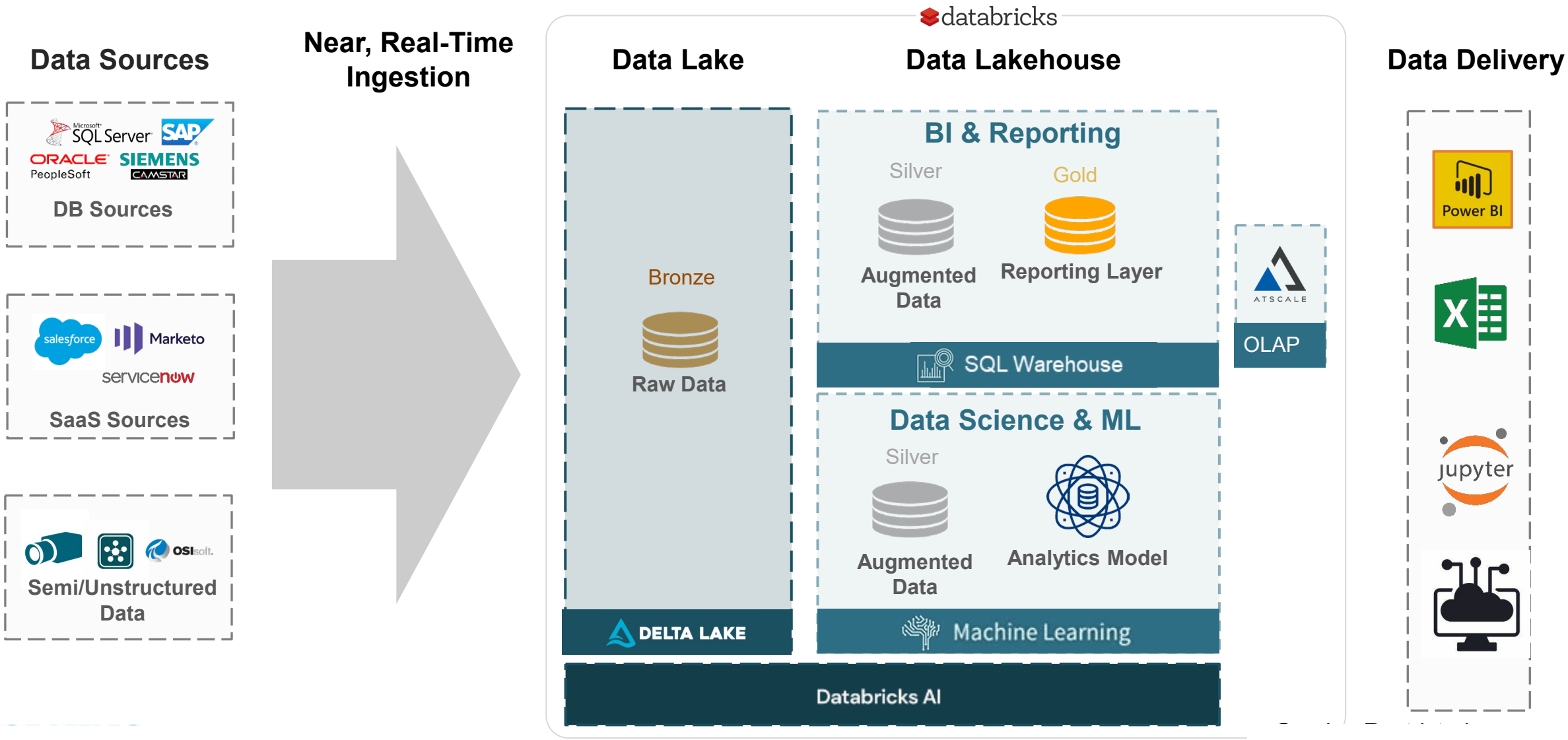
AtScale provides consistent BI & AI Workloads on single platform



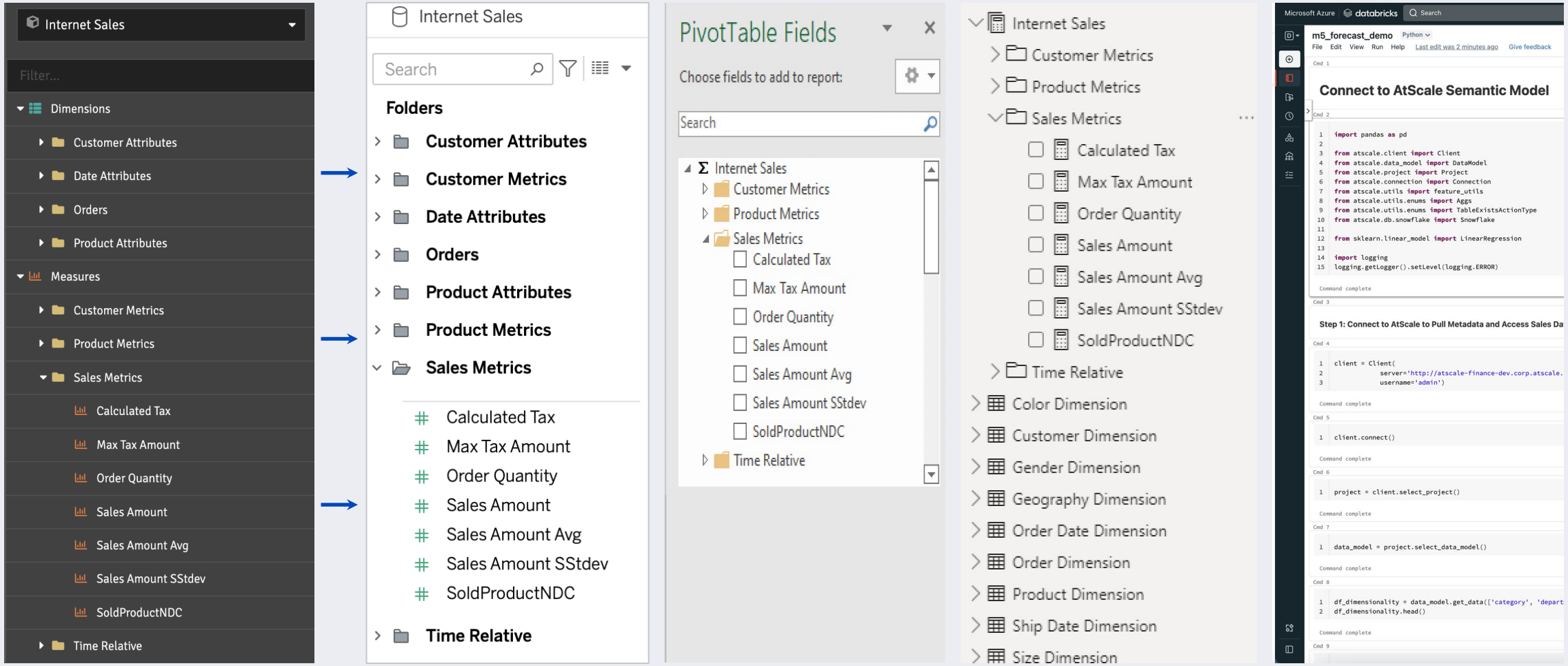

databricks
ATSCALE
Bronze
Raw ingestion and History
Silver
Filtered, Cleaned, Augmented
Gold
Business-level Aggregates
Platinum
AtScale Generated Aggregates
Databricks SQL
Databricks SQL
Databricks SQL
Small Interactive
Large interactive
System Engine
XMLA
JDBC
ODBC
REST
DAX
MDX
SQL
SQL
Python
DELTA LAKE
Microsoft Azure
Google Cloud
aws

Enterprise Data Platform Architecture: Data Processing Framework

This framework adheres to a land once data architecture principle



The Value of a Universal Semantic Layer



AtScale POC & Deployment

Data Scope

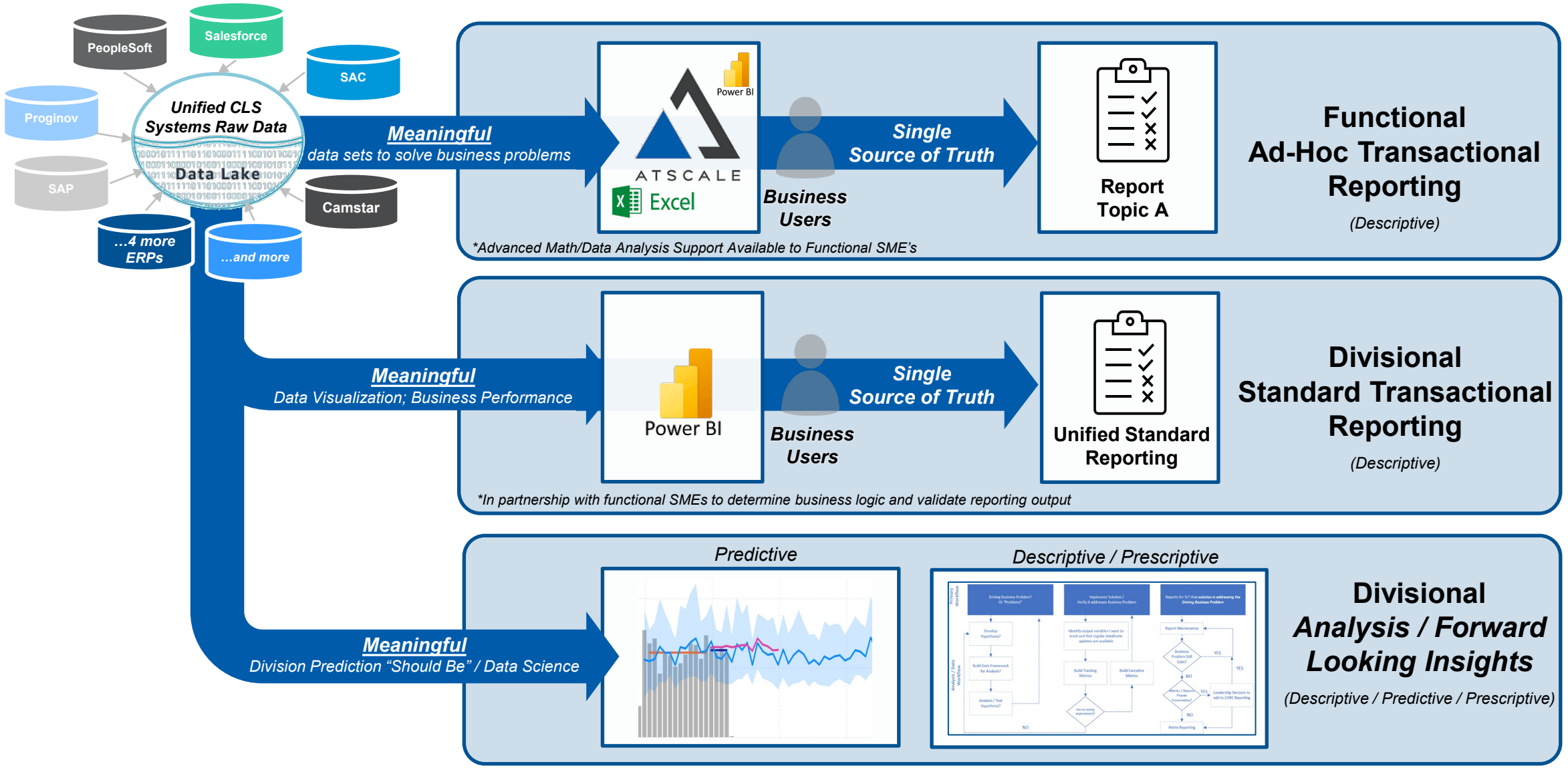
- Built a semantic model of orders and shipment data
- Developed Excel workbooks to connect directly to AtScale semantic
- Deployed the workbooks to 20 power users – the toughest critics

What We Found

- Semantic development was straightforward
- Robust UI for designing metrics and measures
- No need for a calculation period
- User feedback was extremely positive – “It feel like our old data warehouse cube.”

Unified Data; a single source of truth helping functions get to descriptive insights quickly, see single source reporting, and enabling next gen predictive / prescriptive capabilities

BTCOE: Transactional Data Transformation; Enhanced Vision (Set in 2022)



AtScale Next Steps

CLS

- Expand user base of initial model to 100+ users
- Build additional models for customer sales, rebates, contracts, and pricing

CET

- Develop customer, product, and orders semantic in support of first SAP deployment

Finance

- Build a replacement for a manufacturing cost cube

Questions

Thank You!